

Using Time Series Analysis for Evaluating Disease Management Program Effectiveness

Ariel Linden, Dr.P.H., M.S.

Presented by: Ariel Linden, Dr.P.H., M.S., Director of Clinical Quality Improvement, Providence Health Plans, 3601 SW Murray Blvd. Suite 10, Beaverton, OR 97005, Tel: 503-574-6444; Fax: 503-574-6474; E-mail: ariel.linden@providence.org

Research Objective: Currently, the most widely used method in the disease management (DM) industry for evaluating program effectiveness is referred to as the “total population approach.” Effectiveness, in this context, relates to the ability to demonstrate medical cost-savings as a result of the program’s intervention on a specific diseased population. This model is a pretest-posttest design, with the most basic limitation being that without a control group, there may be sources of bias and/or competing extraneous confounding factors that offer plausible rationale explaining the change from baseline. An alternative, and more appropriate approach to evaluating DM program effectiveness is vis-à-vis time series analysis, which is predicated on the notion of serial dependency. In other words, any variable measured over time is influenced by previous observations (autocorrelation). As such, time series models use previous observations as the basis for predicting future behavior. Awareness of temporal influences assist the DM program evaluator identify, describe, explain, and predict the effects of processes that bring about change as a result of the program intervention. This paper proposes a methodology, using time-series analysis, to evaluate DM program effectiveness.

Study Design: 54 months of congestive heart failure (CHF) hospital admission and Emergency Department (ED) data, adjusted per thousand members per year, were used in the analysis. Three time series designs were developed (simple exponential smoothing [SES], double exponential smoothing [DES], and autoregressive integrated moving average [ARIMA]) using (a) the first 36 months data as the historical period, (b) the next 12 months as the validation period, and (c) the final 6 months as the prospective forecast period (representing the first 6 months of a congestive heart DM pilot program). The design eliciting the lowest mean absolute percentage error (MAPE) for the validation period was then used to forecast out to the 6- month prospective period. Program success was determined by comparing actual vs. predicted values in the 6-month period. A mean percentage error (MPE) less than zero would indicate that the program was effective in reducing hospital or ED utilization, whereas a value of zero or greater would indicate no programmatic influence on utilization for the period.

Population Studied: Aggregate data used in this study represents the hospitalization and ED experience of an HMO's CHF population between January 1998 and June 2002.

Principal Findings: CHF Admissions: Following the iterative process of identification, estimation and diagnosis of the time-series data, an ARIMA (1,0,0) was developed. This design proved to be the best fitting model of these data, eliciting a MAPE of 15.5% (compared to 16.0% and 19.7% for SES and DES, respectively). Using the forecasts from this ARIMA model, a MPE of -2.9% was obtained for the prospective period, indicating that the actual admission rate for the period was about 3% lower than predicted. CHF ED Visits: Because this time series appeared to have a trend, the DES proved to best-fit the data of the 3 designs, with a MAPE of 13.9% (compared to 19.8% and 29.2% for SES and ARIMA (1,0,1), respectively). Using the forecasts from the DES model, a MPE of 6.5% was obtained for the prospective period, indicating that the actual ED visit rate was 6.5% higher than predicted.

Conclusions: In the data presented, it appears that the program was effective in reducing hospitalizations for CHF during the 6-month period by 3% over what was predicted. Conversely, it appears that the program was not effective in impacting ED visits, as the actual rate was 6.5% higher than predicted for the period. One important fact to consider when reviewing these results is that DM programs typically do not show an immediate impact on utilization, since the early months of the program are geared toward enrollment and initial patient assessments. If the intervention is effective at the patient level, it may not be evident until several months into the program.

Implications for Policy, Delivery, or Practice: This paper proposed a methodology better suited for evaluating DM program effectiveness than the currently used pretest-posttest design. Time series analysis takes into account the serial dependency of observations in an uncontrolled setting, allowing the DM program evaluator to predict future behavior of the observed variable without attempting to measure independent relationships that influence it. This is an extremely important point, since there are countless factors that may govern the behavior of the time series variable that cannot be identified or accurately measured using the pretest-posttest design.

Primary Funding Source: None.